UCRL-JC-121068 ABS

Genomic sequencing of human chromosome 19 and comparative analysis of human and rodent DNA repair gene regions.

Jane Lamerdin, Mishelle Montgomery, Stephanie Stilwagen, Melissa Ramirez, Aaron Adamson, Subha Basu, Kerry Brookman, Robert Tebbs, Larry Thompson, and Anthony V. Carrano
Human Genome Center, Biology and Biotechnology Research Program, Lawrence Livermore National Laboratory, Livermore, CA, 94550

Approximately 70% of chromosome 19 is spanned by cosmids for which an EcoR1 map has been derived. We are now scaling our sequencing efforts to take advantage of these pre-mapped clones. To date, we have completed ~500 kbp of genomic sequence which has been targeted primarily to cosmids containing the human DNA repair genes *XRCC1* and *ERCC2* on chromosome 19, *ERCC4* on chromosome 16, and *XRCC3* on chromosome 14, as well as selected rodent homologs. We have sequenced 76 kbp containing the human and mouse *XRCC1* genes, which span 26 kbp in the mouse and 31.9 kbp in the human. In addition to the coding regions, 9 conserved elements were identified with sequence identities ranging from 65% to 78%. We have completed 52 kbp of human sequence encompassing the *ERCC2* gene as well as 54 kb spanning the syntenic regions in the mouse and hamster. A defect in *ERCC2* leads to the cancer-prone human disorder xeroderma pigmentosum (XP-D). The human *ERCC2* gene is comprised of 23 exons and is 98% identical to the rodent homologs at the protein level. We identified two genes flanking *ERCC2*, one may be a new member of the kinesin gene family, and the other has no known function. Like *ERCC2*, the *ERCC4* gene product is involved in the nucleotide excision repair pathway, which recognizes and removes DNA damage. A total of 35 kbp has been completed for this gene region which has been instrumental in identifying and assembling the coding regions from numerous partial length cDNAs. The full-length gene spans ~29 kbp and is >50% AT-rich. The *ERCC4* gene product exhibits significant homology to the *S. cerevisiae* rad1 and *S. pombe* rad16 genes, which encode single strand endonucleases. Finally, we have completed sequencing a 2.7 kbp candidate cDNA for the recently cloned human *XRCC3* gene and are in the process of sequencing the cosmid containing this gene, which appears to play a crucial role in chromosomal stability.

This work was performed by Lawrence Livermore National Laboratory under the auspices of the U.S. Department of Energy, Contract No. W-7405-Eng-48.